Cogn-IQ Research Papers Research Article ◆ Open Access doi: pubscience.org/ps-1n0I2-3e4555-nEeG

THE GENDER-EQUALITY PARADOX IN ADULT SKILLS (PIAAC)

The Gender-Equality Paradox in Adult Cognitive Skills: Numeracy, Literacy, and Relative Strengths Across 34 Countries and Two Cycles of PIAAC

Xavier Jouve[1]

[1]Cogn-IQ

Author Note

Xavier Jouve https://orcid.org/0009-0006-1283-045X, Cogn-IQ.

This article analyzes the OECD Programme for the International Assessment of Adult Competencies (PIAAC) public-use files (open data) together with published national indices from the World Economic Forum (Global Gender Gap Index) and the United Nations Development Programme (Human Development Index and Gender Inequality Index). No external funding was received, and all analyses are reproducible from the public data. The author declares no conflict of interest.

Correspondence concerning this article should be addressed to Xavier Jouve, Cogn-IQ. Email: research@cogn-iq.org

ABSTRACT

The gender-equality paradox — the counterintuitive finding that sex differences in academic strengths are larger in more gender-equal countries — is among the most contested results in the study of cognitive sex differences, and it has been established almost entirely in adolescents using PISA. We test whether it holds in adults using the Programme for the International Assessment of Adult Competencies (PIAAC), estimating standardized male-female gaps in numeracy and literacy with plausible values and replicate weights across 34 countries in Cycle 1 (2012) and 24 in Cycle 2 (2023). Three findings emerge. First, adults show a consistent male advantage in numeracy (mean d = 0.20) but essentially no sex gap in literacy (mean d = 0.03), unlike the female reading advantage typical of adolescents. Second, the raw paradox is absent: the numeracy gap is uncorrelated with national gender equality (r = 0.04). Third, the intra-individual paradox — measured as each population's relative numeracy-minus-literacy strength, the metric on which the adolescent paradox rests — does appear and replicates across both independent cycles (r = 0.33 and 0.42 with the WEF Global Gender Gap Index; consistent with the Gender Inequality Index), with country values stable across cycles (r = 0.55). However, controlling for national development (HDI) attenuates the association to r ≈ 0.20-0.24,

indicating that much of the adult paradox is shared with development rather than gender equality per se. The adult paradox is therefore real on the relative-strength measure and reproducible, but weaker than in adolescents and substantially confounded with development.

Keywords: gender-equality paradox, sex differences, numeracy, literacy, PIAAC

1. Introduction

Few empirical claims about cognition attract as much attention, or as much dispute, as those concerning sex differences. The weight of evidence is that average psychological sex differences are small: a synthesis of 46 meta-analyses supported a "gender similarities hypothesis," on which males and females are alike on the large majority of psychological variables, with genuine differences confined to a minority of domains (Hyde, 2005). Cognitive abilities are one of those domains, but even here the differences are modest and domain-specific rather than global. In mathematics, a meta-analysis of 242 studies testing more than 1.2 million people found essentially no overall sex difference ($d = 0.05$) (Lindberg, Hyde, Petersen, & Linn, 2010), and a consensus review concluded that sex differences in science and mathematics ability are small and have multiple, interacting causes (Halpern et al., 2007). What persists is a patterned asymmetry: across large international assessments of adolescents, boys tend to outperform girls in mathematics while girls outperform boys in reading (Stoet & Geary, 2013), and the magnitude of even these differences appears to be changing over historical time (Miller & Halpern, 2014).

The most informative feature of these gaps, for questions of cause, is that they *vary across countries*. A meta-analysis of cross-national mathematics data found that a sex gap persists in some nations but not others, a pattern its authors linked to cultural variation in the status of women — the "gender stratification" hypothesis (Else-Quest, Hyde, & Linn, 2010). Cross-cultural evidence of this kind has been read as implicating social and cultural factors, rather than fixed biological ones, in the size of cognitive sex differences (Reilly, 2012). The natural prediction from a sociocultural account is straightforward: as societies become more gender-equal, sex differences in cognitive performance should shrink.

A prominent body of work reports the opposite. Analyzing international adolescent achievement data, Stoet and Geary described a *gender-equality paradox*: sex differences in academic strengths, and the under-representation of women in science, technology, engineering, and mathematics (STEM), were *larger* in more gender-equal countries (Stoet & Geary, 2018). The finding is genuinely contested. A commentary challenged it on methodological grounds (Richardson et al., 2020), the original authors replied that the paradox is part of a broader phenomenon (Stoet & Geary, 2020), and a recent cross-temporal re-examination affirmed that girls' intra-individual strength is typically reading and boys' is mathematics or science, that these relative strengths are indeed larger in more gender-equal countries, but that the *stability* of the paradox is itself in question (Balducci, 2024).

Two threads in this debate frame the present study. The first concerns *which gap matters*. The paradox is clearest not on raw, single-subject gaps but on an intra-individual contrast — a person's relative strength in one domain over another — because comparative advantage, more than absolute level, channels educational and occupational choice. Girls' comparative advantage in reading can by itself account for much of the gender gap in math-related fields (Breda & Napp, 2019), and relative strengths feature centrally in accounts of why women remain under-represented in STEM despite comparable or superior achievement (Wang & Degol, 2017). The second concerns *scope and confounds*. The paradox has been established almost entirely in *adolescents*, through PISA; whether it characterizes *adults* — whose skills are settled, who are tested by a different program, and who inhabit actual labor markets — is essentially unknown. And because national gender-equality indices are themselves correlated with wealth and human development, a positive gap-equality association may track development rather than gender equality, an objection at the heart of the methodological critique.

The present study. We bring adult data to bear using the Programme for the International Assessment of Adult Competencies (PIAAC), the adult counterpart to PISA, across its two cycles (2012 and 2023) and 34 countries. We ask:

RQ1. What is the profile of adult sex differences in numeracy and literacy, and does it match the adolescent pattern of a male mathematics advantage and a female reading advantage?

RQ2. Is there a *raw* gender-equality paradox in adults — does the single-domain sex gap in numeracy or literacy widen with national gender equality?
RQ3. Does the *intra-individual* paradox — relatively greater male numeracy strength in more gender-equal countries — appear in adults?
RQ4. Does any such paradox *replicate* across two independent PIAAC cycles collected eleven years apart?
RQ5. Is the paradox attributable to gender equality specifically, or is it *confounded with national development*?

2. The Paradox and Its Measurement

2.1 Raw versus intra-individual gaps

Every test of the paradox depends on how the sex gap is defined. A *raw* gap is the standardized male-female difference within a single domain — the numeracy gap or the literacy gap taken alone. An *intra-individual* gap instead captures relative strength: the degree to which one domain exceeds the other within the same person, summarized at the population level as the sex difference in that within-person contrast. The distinction is not cosmetic. Educational and career trajectories respond to comparative advantage — what a person is relatively good at — more than to absolute ability, so the intra-individual contrast is the theoretically motivated quantity, and it is the metric on which the adolescent paradox is strongest (Breda & Napp, 2019); (Balducci, 2024). A paradox that appears on the intra-individual measure but not on raw gaps is therefore both expected and informative, and we report both throughout.

2.2 The confound with development

National gender-equality indices co-vary with economic development: wealthier, higher-development countries tend also to score as more gender-equal. Consequently a cross-country association between a sex gap and gender equality may instead reflect development — through schooling quality, test familiarity, or labor-market structure — rather than gender equality itself. This is the central objection raised against the original

adolescent finding (Richardson et al., 2020). Any credible test of the paradox must therefore ask whether the association survives statistical control for development, not merely whether it exists.

2.3 From adolescents to adults

There are positive reasons to expect the adult picture to differ from the adolescent one. The female reading advantage is large and well documented in 15-year-olds but may narrow with age as schooling effects fade and verbal demands of adult life equalize; the male numeracy advantage, tied to quantitatively intensive education and occupations, may persist or grow. If so, the adult intra-individual contrast would be carried mainly by numeracy, and its relationship to national gender equality need not mirror the adolescent result. Testing the paradox in adults is thus not a redundant replication but a genuine extension to a population in which the constituent gaps may themselves have changed.

3. Data and Measures

3.1 PIAAC and its two cycles

PIAAC is the OECD's international assessment of adult skills, the adult analogue of PISA. We use the public-use files for both available cycles: Cycle 1, fielded in 2012, 2014, or 2017 depending on country (34 countries in our analysis), and Cycle 2, fielded in 2023 (24 of the same countries) (OECD, 2024). Using both cycles provides two independent population samples per country, separated by roughly a decade, and so permits a direct test of replication.

3.2 Sample and domains

Analyses cover adults aged 16-65, the PIAAC target population. Proficiency is carried by the two domains common to both cycles, *numeracy* and *literacy*; the third domain differs between cycles (problem-solving in technology-rich environments in Cycle 1, adaptive problem solving in Cycle 2) and is not comparable, so it is excluded. Each domain is represented by ten plausible values.

3.3 Proficiency: plausible values

Because each respondent is administered only a subset of the item pool, PIAAC does not report a single ability score; it reports, for each respondent and domain, ten *plausible values* — random draws from the posterior distribution of the latent proficiency given the person's responses and a population model (Mislevy, Beaton, Kaplan, & Sheehan, 1992). Correct analysis runs the model on each plausible value and combines the ten by Rubin's rules, so that the variation across them propagates measurement uncertainty into the standard errors. We use the plausible values exactly as published.

3.4 Sex-gap and intra-individual measures

For each country and domain we compute the standardized sex gap (Cohen's d) as the survey-weighted male mean minus the female mean, divided by the pooled standard deviation, with positive values denoting a male advantage. The intra-individual gap is the numeracy d minus the literacy d, which is algebraically the sex difference in within-person standardized relative strength (numeracy relative to literacy). One inferential step deserves to be made explicit. The quantity is a population mean of a within-person contrast, and it is informative about comparative advantage precisely because the contrast is formed within each individual before being averaged: each person contributes the degree to which their own numeracy exceeds their own literacy, so the mean is the average individual-level relative strength, not a difference of two independently averaged domains. What it does not do is describe any particular individual—the within-person contrasts vary widely around their means—so the measure supports claims about the central tendency of comparative advantage in a population, and not about which domain any given person favors. Per-country values are reported in Table 1.

3.5 National indices

Three country-level moderators are used. The WEF Global Gender Gap Index (GGGI) scores gender equality from 0 to 1, higher being more equal (World Economic Forum, 2025). The UNDP Gender Inequality Index (GII) is signed in the opposite direction, higher denoting more inequality, and so should carry the opposite sign — a useful

internal check. The Human Development Index (HDI) indexes national development and serves as the control for the development confound (UNDP, 2024).

4. Method

4.1 Estimating gaps with plausible values and replicate weights

Each country-by-domain gap is estimated on all ten plausible values and combined by Rubin's rules; the sampling variance for each plausible value is obtained from the 80 jackknife replicate weights supplied with the data, and the total variance adds the between-plausible-value (measurement) component. This yields a standardized gap and a standard error per country that respect both the complex sampling design and the imputation of proficiency.

4.2 Correlational design

The paradox is a country-level phenomenon, so the country is the unit of analysis. We correlate each country-level gap (numeracy, literacy, and intra-individual) with each national index across countries, separately by cycle, using Pearson correlations; Spearman correlations yield the same conclusions and are not reported separately.

4.3 Controlling for development

To address RQ5 we compute the partial correlation of the intra-individual gap with the GGGI controlling for HDI. A paradox specific to gender equality should survive this control; one that is largely a development effect should attenuate toward zero.

4.4 Replication and robustness

Replication (RQ4) is assessed two ways: by comparing the Cycle 1 and Cycle 2 correlations, and by correlating the same countries' intra-individual gaps across the two independent cycles. Robustness is examined by restricting the sample to prime-age

adults (25-54), so that the result does not depend on the youngest or oldest cohorts, and by substituting the GII for the GGGI.

4.5 External validation

As a check that the plausible values and weights are handled correctly, we verified that our survey-weighted national proficiency means reproduce the OECD's published country means to within 0.4 score points before computing any gap.

Figure 1 summarizes the workflow.

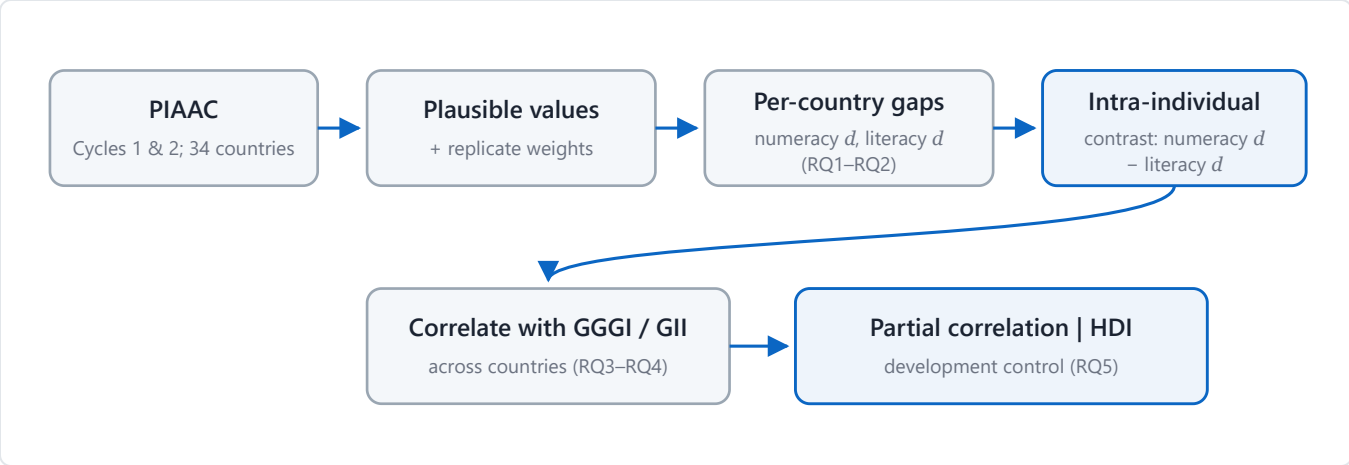

Figure 1. The analysis in one view. Country-level gaps are computed from PIAAC's plausible values with replicate weights; the raw numeracy and literacy gaps answer RQ1–RQ2, their within-person contrast defines the intra-individual measure, and that measure is correlated with the gender-equality indices (RQ3–RQ4) and re-estimated partialling out national development (RQ5).

5. Results

5.1 The adult sex-gap profile (RQ1)

Adults show a consistent male advantage in numeracy and essentially no sex gap in literacy (Figure 2; Table 1). The mean numeracy gap is d = 0.20 in Cycle 1 and d = 0.16 in Cycle 2, positive and statistically distinguishable from zero in nearly every country; the largest is Turkey (d = 0.48), while Russia and Kazakhstan are near zero or slightly female-favoring. The mean literacy gap is d = 0.03 (Cycle 1) and d = -0.05 (Cycle 2) —

close to zero, and conspicuously *not* the female reading advantage characteristic of 15-year-olds. The adult profile is thus asymmetric: a robust male numeracy advantage paired with literacy parity. This already answers RQ1 in the negative for one half of the adolescent pattern — the female literacy advantage does not survive into adulthood in these data.

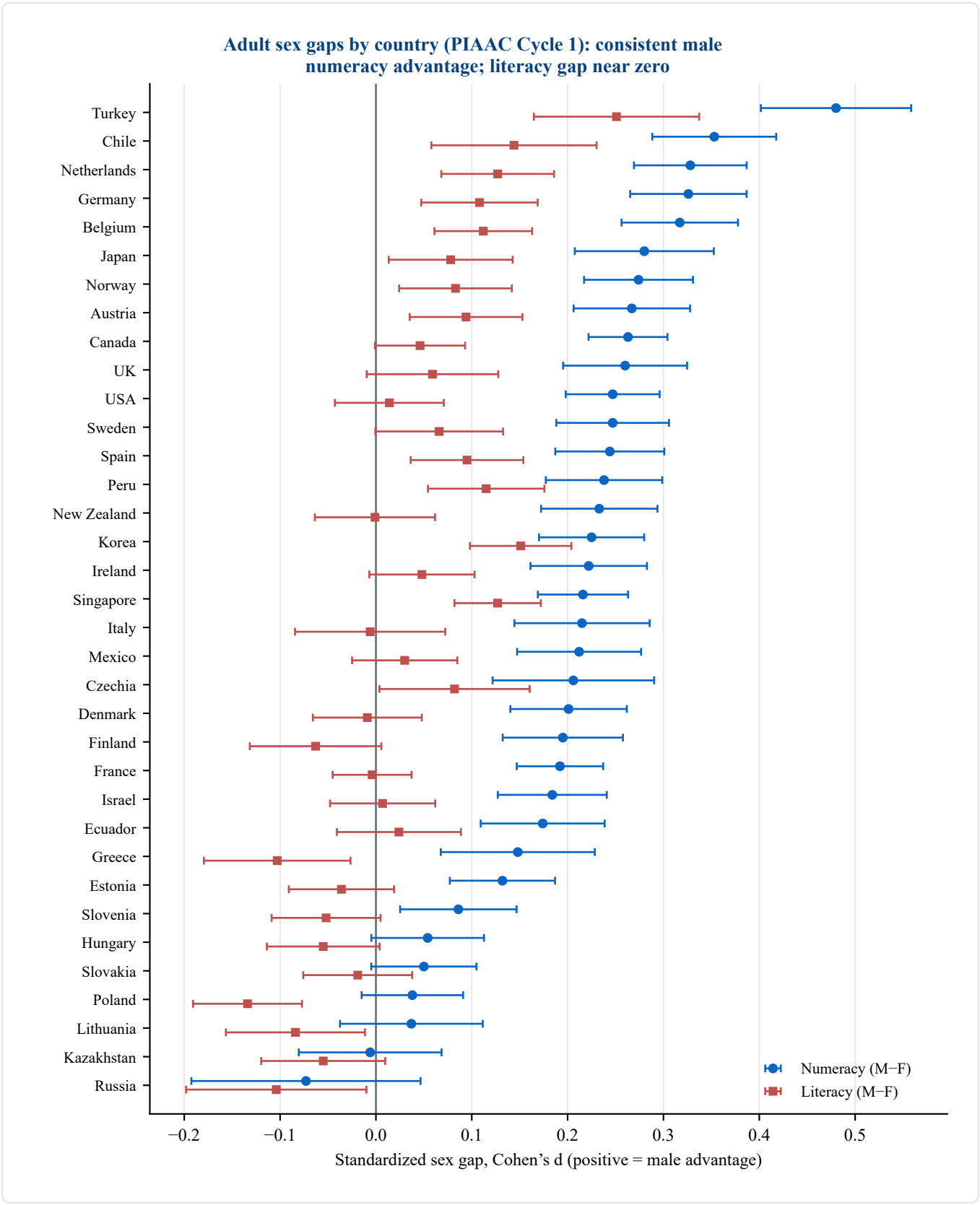

Figure 2. Standardized adult sex gaps by country (PIAAC Cycle 1), with 95% confidence intervals. The numeracy gap (male advantage) is positive almost everywhere; the literacy gap clusters near zero.

Table 1

Standardized adult sex gaps (Cohen's d, male − female) in numeracy and literacy, with the intra-

individual gap, by country and cycle. Countries ordered by gender equality

Country	GGGI	HDI	C1 d_{num}	C1 d_{lit}	C1 intra	C2 d_{num}	C2 d_{lit}	C2 intra
Finland	0.88	0.95	+0.20 ±0.03	-0.06 ±0.04	+0.26	+0.17 ±0.19	-0.11 ±0.19	+0.29
Norway	0.86	0.97	+0.27 ±0.03	+0.08 ±0.03	+0.19	+0.18 ±0.23	-0.07 ±0.22	+0.26
United Kingdom	0.84	0.94	+0.26 ±0.03	+0.06 ±0.04	+0.20	+0.27 ±0.23	+0.04 ±0.23	+0.23
New Zealand	0.83	0.94	+0.23 ±0.03	-0.00 ±0.03	+0.23	+0.08 ±0.39	-0.20 ±0.38	+0.28
Sweden	0.82	0.96	+0.25 ±0.03	+0.07 ±0.03	+0.18	+0.23 ±0.24	-0.06 ±0.27	+0.29
Germany	0.80	0.96	+0.33 ±0.03	+0.11 ±0.03	+0.22	+0.20 ±0.19	-0.07 ±0.19	+0.28
Ireland	0.80	0.95	+0.22 ±0.03	+0.05 ±0.03	+0.17	+0.18 ±0.26	-0.07 ±0.26	+0.26
Estonia	0.80	0.91	+0.13 ±0.03	-0.04 ±0.03	+0.17	+0.13 ±0.15	-0.20 ±0.14	+0.33
Spain	0.80	0.92	+0.24 ±0.03	+0.10 ±0.03	+0.15	+0.19 ±0.19	+0.04 ±0.19	+0.15
Denmark	0.79	0.96	+0.20 ±0.03	-0.01 ±0.03	+0.21	+0.18 ±0.22	-0.06 ±0.22	+0.24
Lithuania	0.78	0.90	+0.04 ±0.04	-0.08 ±0.04	+0.12	+0.09 ±0.20	-0.08 ±0.21	+0.17
Chile	0.78	0.88	+0.35 ±0.03	+0.14 ±0.04	+0.21	+0.14 ±0.17	-0.01 ±0.16	+0.15
Mexico	0.78	0.79	+0.21 ±0.03	+0.03 ±0.03	+0.18	—	—	—
Ecuador	0.77	0.78	+0.17 ±0.03	+0.02 ±0.03	+0.15	—	—	—
Belgium	0.77	0.95	+0.32 ±0.03	+0.11 ±0.03	+0.20	+0.19 ±0.22	-0.01 ±0.21	+0.20

Country	GGGI	HDI	C1 d_{num}	C1 d_{lit}	C1 intra	C2 d_{num}	C2 d_{lit}	C2 intra
Slovenia	0.77	0.93	+0.09 ±0.03	-0.05 ±0.03	+0.14	—	—	—
Canada	0.77	0.94	+0.26 ±0.02	+0.05 ±0.02	+0.22	+0.28 ±0.20	+0.01 ±0.21	+0.27
France	0.77	0.92	+0.19 ±0.02	-0.00 ±0.02	+0.20	+0.17 ±0.03	-0.06 ±0.03	+0.23
Netherlands	0.76	0.95	+0.33 ±0.03	+0.13 ±0.03	+0.20	—	—	—
United States	0.76	0.94	+0.25 ±0.03	+0.01 ±0.03	+0.23	+0.12 ±0.15	-0.05 ±0.16	+0.17
Poland	0.75	0.91	+0.04 ±0.03	-0.13 ±0.03	+0.17	+0.00 ±0.19	-0.06 ±0.17	+0.06
Singapore	0.75	0.95	+0.22 ±0.02	+0.13 ±0.02	+0.09	+0.17 ±0.19	+0.07 ±0.17	+0.10
Austria	0.73	0.93	+0.27 ±0.03	+0.09 ±0.03	+0.17	+0.25 ±0.18	-0.00 ±0.20	+0.26
Peru	0.73	0.79	+0.24 ±0.03	+0.12 ±0.03	+0.12	—	—	—
Slovakia	0.72	0.88	+0.05 ±0.03	-0.02 ±0.03	+0.07	+0.03 ±0.20	-0.01 ±0.19	+0.04
Greece	0.71	0.91	+0.15 ±0.04	-0.10 ±0.04	+0.25	—	—	—
Israel	0.71	0.92	+0.18 ±0.03	+0.01 ±0.03	+0.18	—	—	—
Italy	0.70	0.92	+0.21 ±0.04	-0.01 ±0.04	+0.22	+0.12 ±0.20	-0.06 ±0.20	+0.18
Kazakhstan	0.70	0.84	-0.01 ±0.04	-0.06 ±0.03	+0.05	—	—	—
Korea	0.69	0.94	+0.23 ±0.03	+0.15 ±0.03	+0.07	—	—	—
Czechia	0.69	0.92	+0.21 ±0.04	+0.08 ±0.04	+0.12	+0.19 ±0.23	-0.03 ±0.24	+0.22

Country	GGGI	HDI	C1 d_{num}	C1 d_{lit}	C1 intra	C2 d_{num}	C2 d_{lit}	C2 intra
Hungary	0.68	0.87	+0.05 ±0.03	-0.06 ±0.03	+0.11	+0.08 ±0.20	-0.12 ±0.20	+0.20
Japan	0.67	0.93	+0.28 ±0.04	+0.08 ±0.03	+0.20	+0.23 ±0.20	-0.00 ±0.20	+0.23
Turkey	0.63	0.85	+0.48 ±0.04	+0.25 ±0.04	+0.23	—	—	—

Note. d ± jackknife SE. C1 = Cycle 1 (2012/14/17); C2 = Cycle 2 (2023); — = country not in Cycle 2. Intra = $d_{num} - d_{lit}$. GGGI = WEF Global Gender Gap Index (higher = more equal); HDI = Human Development Index. Source: PIAAC public-use files; WEF; UNDP.

5.2 No raw paradox (RQ2)

On raw, single-domain gaps there is no adult gender-equality paradox (Table 2). The numeracy gap is uncorrelated with national gender equality: $r = 0.04$ ($p = 0.84$) in Cycle 1 and $r = 0.19$ ($p = 0.37$) in Cycle 2. The literacy gap is likewise unrelated to equality. Whatever the adolescent literature shows on raw subject scores, the simple version of the paradox is absent in adult numeracy and literacy.

5.3 The intra-individual paradox (RQ3)

On the intra-individual relative-strength measure the relationship emerges (Figure 3). Relative male numeracy strength rises with national gender equality: $r = 0.33$ ($p = 0.055$) in Cycle 1 and $r = 0.42$ ($p = 0.040$) in Cycle 2. The evidence is consistent across cycles, although the Cycle 1 association narrowly misses conventional significance; the two cycles are independent samples, and it is their agreement rather than either p value that carries the weight here. The Gender Inequality Index, signed in the opposite direction, gives correlations of the expected opposite sign (-0.20 and -0.46), confirming the pattern is not an artifact of a single index. RQ3 is therefore answered in the affirmative, but with the immediate qualification that the effect is moderate and, in Cycle 1, only borderline by conventional thresholds.

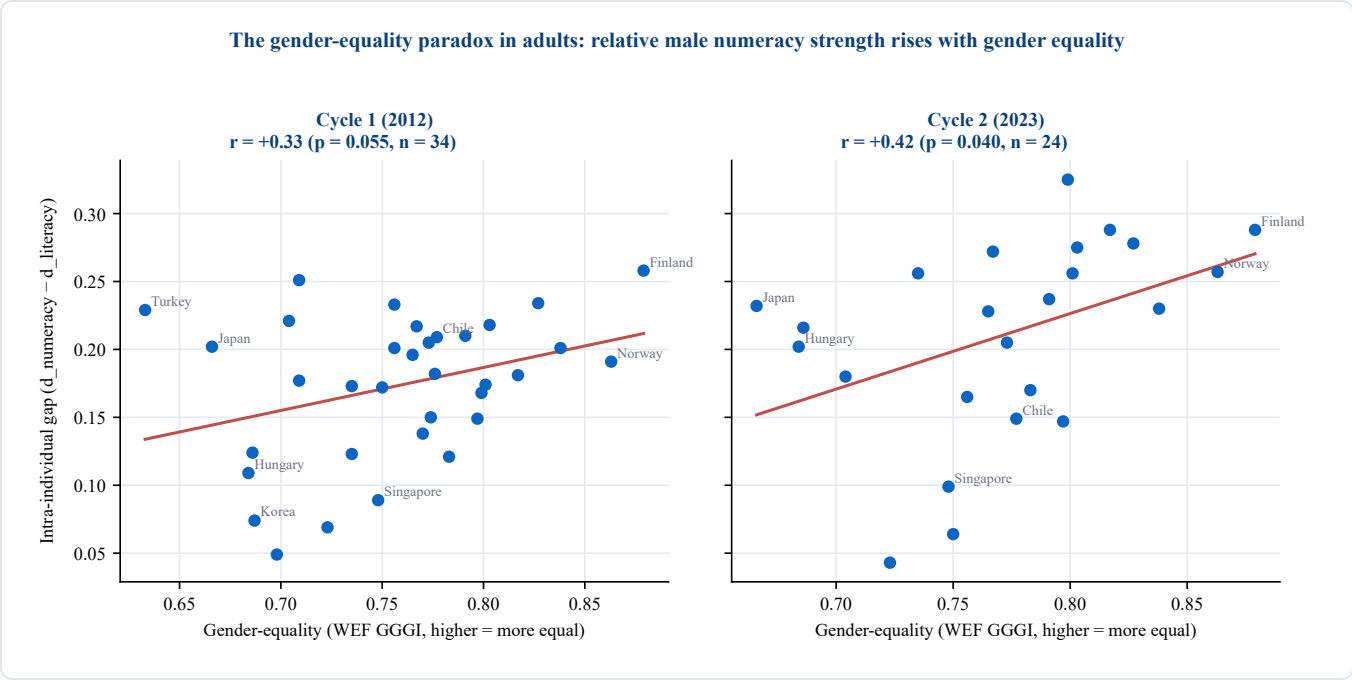

Figure 3. The intra-individual gap (numeracy − literacy) versus national gender equality, both cycles. The positive slope — relatively greater male numeracy strength in more gender-equal countries — is the gender-equality paradox, here in adults. Turkey and Japan are notable counter-cases (low equality, high gap).

5.4 Replication across cycles (RQ4)

The intra-individual paradox is not a one-sample artifact. Both cycles, drawn from independent samples eleven years apart, show the same positive association (0.33 and 0.42), and the countries' intra-individual gaps themselves correlate $r = 0.55$ between cycles (Figure 5). The cross-national pattern of relative strengths is thus a reproducible property of populations rather than sampling noise — a point that bears directly on the debated *stability* of the paradox (Balducci, 2024). RQ4 is answered affirmatively.

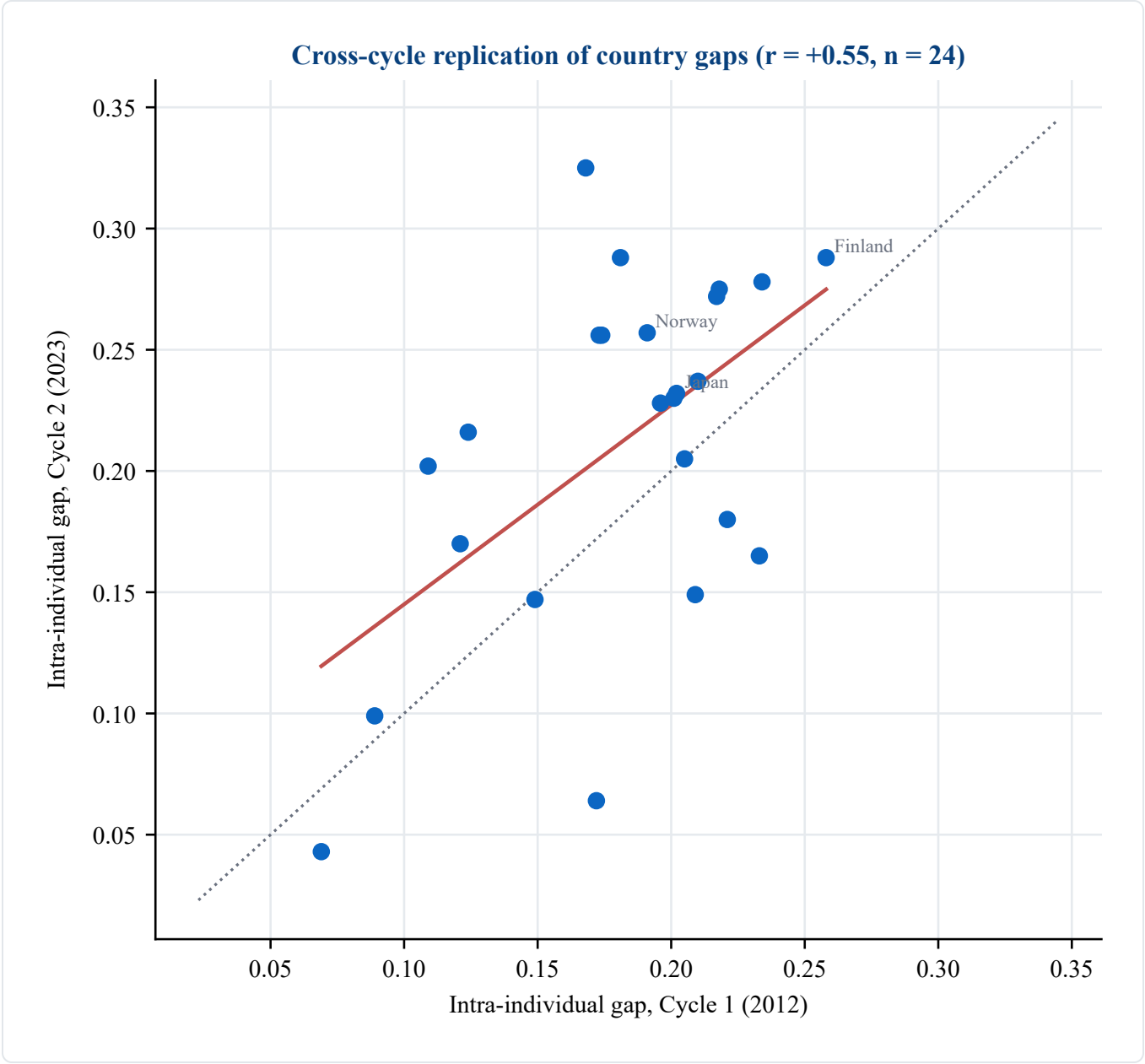

Figure 5. Cross-cycle replication: countries' intra-individual gaps are stable between two independent samples (2012 and 2023), r = 0.55.

5.5 Robustness: indices and age

The association is robust to the analytic choices most likely to matter. Restricting to prime-age adults (25-54) — removing the youngest, still-in-education cohorts and the oldest, partly-retired ones — leaves the Cycle 1 correlation essentially unchanged at r = 0.30 (p = 0.083). Substituting the GII for the GGGI preserves the pattern with the expected sign reversal. The intra-individual paradox is therefore not an artifact of age composition or of one particular equality index.

5.6 The development confound (RQ5)

The decisive test is whether the paradox is specific to gender equality or shared with development (Figure 4; Table 2). It is substantially shared. Gender equality and development are themselves correlated across these countries (GGGI-HDI r = 0.36), and partialling out HDI attenuates the intra-individual association from 0.33 to 0.24 in Cycle 1 and from 0.42 to 0.20 in Cycle 2. Roughly half of the apparent paradox is thus attributable to development rather than to gender equality per se. Because GGGI and HDI are themselves correlated, the partial correlations should be read with the usual caution: with collinear predictors, the split of shared variance between them is not uniquely determined, and the partial estimate is best understood as the association that survives when the development component is removed rather than as a clean causal decomposition. Two countries sharpen the point: Turkey and Japan combine low gender equality with large intra-individual gaps, sitting clearly against the paradox and reminding that the cross-national relationship is a modest central tendency, not a deterministic law. RQ5 is therefore answered with a qualified negative: the paradox is not cleanly attributable to gender equality alone.

Table 2

Cross-country correlations of adult sex gaps with national gender-equality and development indices, by cycle

Sex-gap measure	GGGI C1	GGGI C2	GII C1	GII C2	GGGI·HDI C1	GGGI·HDI C2
Numeracy gap (M−F)	+0.04	+0.19	-0.01	-0.48	-0.05	-0.20
Literacy gap (M−F)	-0.17	-0.26	+0.11	-0.01	-0.20	-0.37
Intra-individual (num − lit)	+0.33	+0.42	-0.20	-0.46	+0.24	+0.20

Note. Entries are cross-country Pearson correlations. GGGI = WEF Global Gender Gap Index (higher = more equal); GII = UNDP Gender Inequality Index (higher = more unequal; expected opposite sign). GGGI·HDI = partial correlation with GGGI controlling for the Human Development Index. C1 n = 34; C2 n = 24.

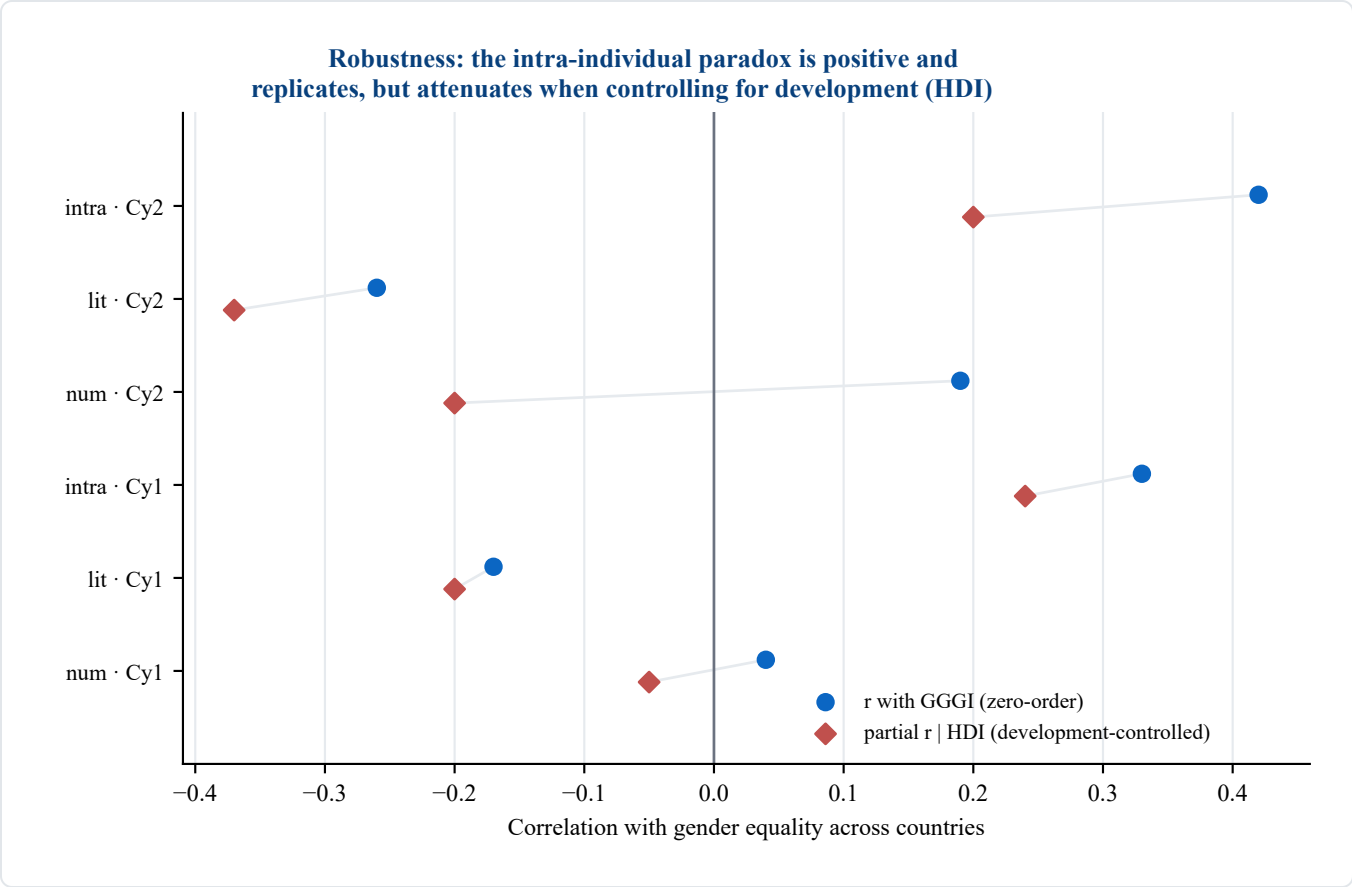

Figure 4. Zero-order versus development-controlled (partial-HDI) correlations of each gap measure with gender equality. The intra-individual paradox is positive and replicates, but shrinks when development is held constant.

6. Discussion

The five questions resolve as follows. **RQ1.** Adults match the adolescent pattern only in part — the male numeracy advantage is present (mean d = 0.20) but the female literacy advantage is not (mean d ≈ 0). **RQ2.** There is no raw paradox: the single-domain numeracy gap is uncorrelated with national gender equality. **RQ3.** The intra-individual paradox does appear (r = 0.33 and 0.42). **RQ4.** It replicates across both independent cycles, with stable country values (r = 0.55). **RQ5.** It is not cleanly attributable to gender equality — controlling for development roughly halves it (partial r ≈ 0.20–0.24). The sections below develop each point.

6.1 A metric-dependent paradox

The adult evidence reproduces, in a new population, the single most important methodological lesson of the adolescent debate: whether a gender-equality paradox

appears depends on how the sex gap is defined. On raw numeracy or literacy gaps there is nothing (RQ2); on the intra-individual relative-strength contrast there is a positive, replicable association (RQ3, RQ4). This is direct adult support for the view that the paradox is a property of *relative* strengths rather than of any single ability (Richardson et al., 2020); (Balducci, 2024), consistent with the comparative-advantage logic that links relative strengths to field choice (Breda & Napp, 2019).

6.2 Entanglement with development

On the measure where it appears, the adult paradox is real and reproducible but modest, and roughly half of it is shared with national development. Controlling HDI leaves an association of only r ≈ 0.20-0.24, and gender equality and development are too correlated in these data to be cleanly separated. The honest reading is that gender equality and development jointly track relative cognitive strengths and that the present design — like the cross-national designs that preceded it — cannot assign the effect to gender equality alone. This is the empirical substance behind the contested status of the paradox rather than a verdict for either side.

6.3 The disappearance of the adult literacy advantage

A striking and, to our knowledge, under-emphasized result is that the female reading advantage so prominent in adolescents is absent in adults: the mean adult literacy gap is essentially zero, while the male numeracy advantage persists at d ≈ 0.20. The adult intra-individual contrast is therefore carried almost entirely by the numeracy side. Because cognitive sex differences are known to change across development and historical time (Miller & Halpern, 2014), this adolescent-to-adult shift is plausible, but it means the adult paradox is structurally different from the adolescent one even where the cross-national pattern looks similar.

6.4 Adults versus adolescents

Set against the adolescent literature — where the raw math gap favors boys and the raw reading gap favors girls across most of 75 countries (Stoet & Geary, 2013), and where overall math differences are negligible in magnitude (Lindberg et al., 2010) — the adult

data agree on a persistent numeracy gap but diverge on literacy. That the intra-individual paradox nonetheless appears in adults extends the phenomenon beyond the school years and the PISA program on which it was built, while its attenuation under development control qualifies how much explanatory weight it can carry.

6.5 Implications

For interpreting national skill statistics and STEM-participation debates, the practical implication is twofold. First, comparisons that invoke "the gender gap" should specify whether they mean a raw or a relative-strength gap, because the two behave differently. Second, claims that greater gender equality *widens* cognitive sex differences should, for adults, be made only with the development confound made explicit, since much of the cross-national signal is shared with development. The comparative-advantage channel that links relative strengths to occupational choice (Wang & Degol, 2017) may well operate in adults, but the present results locate it in a modest, development-entangled cross-national pattern rather than a strong effect of gender equality alone.

7. Limitations

Several limitations bound these conclusions. The paradox analysis is ecological: correlations are computed across 24–34 countries, and country-level ecological analyses necessarily have modest statistical power — with 34 countries, a correlation must reach about .34 to be conventionally significant, so estimates of this size are inherently unstable, the Cycle 1 result is borderline, and estimates are sensitive to individual countries (notably Turkey and Japan) and to index choice — risks we mitigate with two indices, two cycles, and a development control, but cannot eliminate. The design is repeated cross-sections rather than a panel, so age and cohort effects are not separated. The intra-individual measure is itself the subject of ongoing methodological dispute (Richardson et al., 2020). The gender-equality indices are not perfectly period-matched to each cycle's collection year, although their cross-national rank order is highly stable over the relevant span. Finally, PIAAC carries two comparable domains rather than the three available in adolescent assessments, so the relative-strength contrast is numeracy-versus-

literacy only, and the development control rests on a single composite index (HDI) rather than a full model of national context.

8. Conclusion

In adults, the gender-equality paradox appears only on the intra-individual relative-strength measure, not on raw subject gaps; it replicates across two independent PIAAC cycles with stable country values; but it is weaker than in adolescents and substantially confounded with national development. Adult literacy shows no female advantage, so the adult version of the paradox is carried by a persistent male numeracy advantage rather than by a balance of opposing gaps. The adult data thus both *extend* the paradox beyond the adolescent, school-based evidence on which it was built and *qualify* it — as metric-dependent, development-entangled, and modest in size. Claims that rising gender equality widens cognitive sex differences should, for adults, be advanced only with those qualifications attached.

Figures & tables

Figure 1. *The analysis in one view. Country-level gaps are computed from PIAAC's plausible values with replicate weights; the raw numeracy and literacy gaps answer RQ1–RQ2, their within-person contrast defines the intra-individual measure, and that measure is correlated with the gender-equality indices (RQ3–RQ4) and re-estimated partialling out national development (RQ5).*

Figure 2. *Standardized adult sex gaps by country (PIAAC Cycle 1), with 95% confidence intervals. The numeracy gap (male advantage) is positive almost everywhere; the literacy gap clusters near zero.*

Figure 3. *The intra-individual gap (numeracy − literacy) versus national gender equality, both cycles. The positive slope — relatively greater male numeracy strength in more gender-equal countries — is the gender-equality paradox, here in adults. Turkey and Japan are notable counter-cases (low equality, high gap).*

Figure 4. *Zero-order versus development-controlled (partial-HDI) correlations of each gap measure with gender equality. The intra-individual paradox is positive and replicates, but shrinks when development is held constant.*

Figure 5. *Cross-cycle replication: countries' intra-individual gaps are stable between two independent samples (2012 and 2023), r = 0.55.*

Table 1. *Standardized adult sex gaps (Cohen's d, male − female) in numeracy and literacy, with the intra-individual gap, by country and cycle. Countries ordered by gender equality*

Table 2. *Cross-country correlations of adult sex gaps with national gender-equality and development indices, by cycle*

References

Balducci, M. (2024). The gender-equality paradox in intraindividual academic strengths: A cross-temporal analysis. *Psychological Science.* https://doi.org/10.1177/09567976241271330

Breda, T., & Napp, C. (2019). Girls' comparative advantage in reading can largely explain the gender gap in math-related fields. *Proceedings of the National Academy of Sciences.* https://doi.org/10.1073/pnas.1905779116

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin.* https://doi.org/10.1037/a0018053

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest.* https://doi.org/10.1111/j.1529-1006.2007.00032.x

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist.* https://doi.org/10.1037/0003-066X.60.6.581

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin.* https://doi.org/10.1037/a0021276

Miller, D. I., & Halpern, D. F. (2014). The new science of cognitive sex differences. *Trends in Cognitive Sciences.* https://doi.org/10.1016/j.tics.2013.10.011

Mislevy, R. J., Beaton, A. E., Kaplan, B., & Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement.* https://doi.org/10.1111/j.1745-3984.1992.tb00371.x

OECD (2024). *Do Adults Have the Skills They Need to Thrive in a Changing World? Survey of Adult Skills 2023.* OECD Publishing.

Reilly, D. (2012). Gender, culture, and sex-typed cognitive abilities. *PLoS ONE.* https://doi.org/10.1371/journal.pone.0039904

Richardson, S., Sanchez Manzano, F., et al. (2020). Is there a gender-equality paradox in science, technology, engineering, and math (STEM)? Commentary on the study by Stoet and Geary (2018). *Psychological Science.* https://doi.org/10.1177/0956797619872762

Stoet, G., & Geary, D. C. (2018). The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychological Science.* https://doi.org/10.1177/0956797617741719

Stoet, G., & Geary, D. C. (2020). The gender-equality paradox is part of a bigger phenomenon: Reply to Richardson and colleagues. *Psychological Science.* https://doi.org/10.1177/0956797620904134

Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within- and across-nation assessment of 10 years of PISA data. *PLoS ONE.* https://doi.org/10.1371/journal.pone.0057988

United Nations Development Programme (2024). *Human Development Report 2023-24.* UNDP.

Wang, M.-T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review.* https://doi.org/10.1007/s10648-015-9355-x

World Economic Forum (2025). *Global Gender Gap Report 2025.* World Economic Forum.

CITE THIS ARTICLE

Jouve, X. (2026). The gender-equality paradox in adult cognitive skills: Numeracy, literacy, and relative strengths across 34 countries and two cycles of PIAAC. *Cogn-IQ Research Papers.* https://pubscience.org/ps-1n0I2-3e4555-nEeG

Export: BibTeX · RIS · EndNote · APA · Plain text

OUTLINE

Abstract

1. Introduction

2. The Measurement Problem

3. Data & Measures
4. Method
5. Results
6. Discussion
7. Limitations
8. Conclusion
References

ARTICLE

JOURNAL
Cogn-IQ Research Papers

TYPE
Research Article · Intelligence & Cognition

PUBLISHED
June 9, 2026

ACCESS
Open Access

DOI
pubscience.org/ps-1n0I2-3e4555-nEeG

TOOLS

- Full text
- Figures & tables
- Cite this article
- References

CONTINUE ON COGN-IQ

Related Educational Resources

- The g factor — the general ability numeracy and literacy both reflect.
- The AFQT as a measure of general cognitive ability — a companion large-scale cognitive-skill analysis.